

**Manuscript version: Author's Accepted Manuscript**

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

**Persistent WRAP URL:**

<http://wrap.warwick.ac.uk/138619>

**How to cite:**

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk).

# Lyman continuum leakage in faint star-forming galaxies at redshift $z = 3 - 3.5$ probed by gamma-ray bursts <sup>★</sup>

J.-B. Vielhaure<sup>1★★</sup>, S. D. Vergani<sup>1,2</sup>, J. Japelj<sup>3</sup>, J. P. U. Fynbo<sup>4,5</sup>, M. Gronke<sup>6★★★</sup>, K. E. Heintz<sup>7</sup>, D. B. Malesani<sup>4,5,8</sup>, P. Petitjean<sup>2</sup>, N. R. Tanvir<sup>9</sup>, V. D’Elia<sup>10,11</sup>, D. A. Kann<sup>12</sup>, J. T. Palmerio<sup>1</sup>, R. Salvaterra<sup>13</sup>, K. Wiersema<sup>14,15</sup>, M. Arabsalmani<sup>16,17</sup>, S. Campana<sup>18</sup>, S. Covino<sup>18</sup>, M. De Pasquale<sup>19</sup>, A. de Ugarte Postigo<sup>12</sup>, F. Hammer<sup>1</sup>, D. H. Hartmann<sup>20</sup>, P. Jakobsson<sup>7</sup>, C. Kouveliotou<sup>21,22</sup>, T. Laskar<sup>23</sup>, A. J. Levan<sup>15,24</sup> and A. Rossi<sup>25</sup>

(Affiliations can be found after the references)

Received ... ; accepted ...

## ABSTRACT

**Context.** The identification of the sources that reionized the Universe and their specific contribution to this process are key missing pieces of our knowledge of the early Universe. Faint star-forming galaxies may be the main contributors to the ionizing photon budget during the epoch of reionization (EoR), but their escaping photons cannot be detected directly due to inter-galactic medium opacity. Hence, it is essential to characterize the properties of faint galaxies with significant Lyman continuum (LyC) photon leakage up to  $z \sim 4$  to define indirect indicators allowing analogues to be found at the highest redshift.

**Aims.** Long gamma-ray bursts (LGRB) explode typically in star-forming regions of faint, star-forming galaxies. Through LGRB afterglow spectroscopy it is possible to detect directly LyC photons. Our aim is to use LGRBs as tools to study LyC leakage from faint, star-forming galaxies at high redshift.

**Methods.** Here we present the observations of LyC emission in the afterglow spectra of GRB 191004B at  $z = 3.5055$ , together with those of the other two previously known LyC-emitting LGRB (GRB 050908 at  $z = 3.3467$ , and GRB 060607A at  $z = 3.0749$ ), to determine their LyC escape fraction and compare their properties.

**Results.** From the afterglow spectrum of GRB 191004B we determine a neutral hydrogen column density at the LGRB redshift of  $\log(N_{\text{H I}}/\text{cm}^{-2}) = 17.2 \pm 0.15$ , and negligible extinction ( $A_V = 0.03 \pm 0.02$  mag). The only metal absorption lines detected are C IV and Si IV. In contrast to GRB 050908 and GRB 060607A, the host galaxy of GRB 191004B displays significant Ly $\alpha$  emission. From its Ly $\alpha$  emission and the non-detection of Balmer emission lines we constrain its star-formation rate (SFR) to  $1 \leq \text{SFR} \leq 4.7 \text{ M}_{\odot} \text{ yr}^{-1}$ . We fit the Ly $\alpha$  emission with a *shell* model and find parameters values consistent with the observed ones. The absolute (relative) LyC escape fractions we find for GRB 191004B, GRB 050908 and GRB 060607A are of  $0.35^{+0.10}_{-0.11}$  ( $0.43^{+0.12}_{-0.13}$ ),  $0.08^{+0.05}_{-0.04}$  ( $0.08^{+0.05}_{-0.04}$ ) and  $0.20^{+0.05}_{-0.05}$  ( $0.45^{+0.15}_{-0.15}$ ), respectively. We compare the LyC escape fraction of LGRBs to the values of other LyC emitters found from the literature, showing that LGRB afterglows can be powerful tools to study LyC escape for faint high-redshift star-forming galaxies. Indeed we could push LyC leakage studies to much higher absolute magnitudes. The host galaxies of the three LGRB presented here have all  $M_{1600} > -19.5$  mag, with the GRB 060607A host at  $M_{1600} > -16$  mag. LGRB hosts may therefore be particularly suitable for exploring the ionizing escape fraction in galaxies that are too faint or distant for conventional techniques. Furthermore the time investment is very small compared to galaxy studies.

**Key words.** Gamma-ray burst: general, individual: GRB050908, GRB060607A, GRB191004B - Galaxies: high-redshift, evolution, intergalactic medium - dark ages, reionization, first stars

## 1. Introduction

Understanding the nature of the sources responsible for ionizing hydrogen in the intergalactic medium (IGM) remains one of the key challenges in studies of early structure formation. It has been established that active galactic nuclei by themselves provide enough ionizing radiation to keep the Universe reionized at redshifts  $z < 3 - 4$  (e.g., [Cristiani et al. 2016](#)). According to the predominant view, star-forming galaxies (SFGs) are the main contributors of ionizing radiation at earlier epochs ( $z \gtrsim 4$ ; e.g., [Fontanot et al. 2012](#); [Robertson et al. 2013, 2015](#)).

In order to quantify the contribution of SFGs to the reionization at  $z \gtrsim 4$ , we have to know: (i) the galaxy number density as a function of redshift and luminosity; (ii) their ioniz-

ing photon production efficiency (i.e., the number of hydrogen-ionizing photons relative to produced UV photons at  $\sim 1500 \text{ \AA}$ ,  $(f_{1500}/f_{900})^{\text{int}}$ ); (iii) the fraction of produced Lyman continuum (LyC) photons that can actually escape the local environment and ionize the IGM, called the escape fraction  $f_{\text{esc}}$ . Especially the latter is poorly constrained from observations.

Direct searches for LyC emission from galaxies is only possible up to  $z \sim 4$ , beyond which the LyC is unobservable due to the increasing IGM opacity (e.g., [Madau 1995](#)). We therefore rely on lower redshift galaxies to be used as proxies for the higher-redshift ones. The search for LyC emitters has been quite fruitful in recent years. A few LyC emitters have been found in the local Universe ([Bergvall et al. 2006](#); [Leitet et al. 2013](#); [Borthakur et al. 2014](#)). A population of green pea galaxies with strong LyC has been identified at  $z \sim 0.3$  ([Izotov et al. 2016, 2018a,b](#)). A number of LyC-emitting galaxies with typically  $M_{\text{UV}} < -19.5$  mag have also been securely identified at higher ( $z \gtrsim 3$ ) redshifts ([Vanzella et al. 2016, 2018](#); [Shapley et al. 2016](#); [Bian et al. 2017](#); [Steidel et al. 2018](#); [Fletcher et al. 2019](#)). The growing number of discov-

<sup>★</sup> Results based on observations carried out at ESO Observatory, Paranal, Chile, by the Stargate consortium under Program ID: 0104.D-0600, P.I.: N. Tanvir

<sup>★★</sup> e-mail: jean-baptiste.vielhaure@obspm.fr

<sup>★★★</sup> Hubble fellow

ered LyC emitters allows us to look at their common properties that could be used both to make it easier to find new emitters at these redshifts as well as to identify likely emitters at  $z > 6$ . The emitters are typically (but not always) found to have a strong Ly $\alpha$  emission line (Verhamme et al. 2015, 2017), a high ratio of [OIII]/[OII] nebular emission lines (Nakajima & Ouchi 2014; Nakajima et al. 2016) and a compact morphology (Izotov et al. 2016, 2018a,b). The compactness of the galaxy as a requirement for LyC photons to escape is also found in theoretical studies Cen (e.g., 2020), together with high star-formation rate surface densities.

Recently, important constraints on the LyC emission at high redshift come from an analysis of stacked spectra of galaxies at  $-19.5 > M_{UV} > -22$  mag at  $z \sim 3$  (Steidel et al. 2018). The measured mean escape fraction of  $0.09 \pm 0.01$  is dominated by sub- $L_{UV}^*$  galaxies, with negligible contribution from  $L > L_{UV}^*$  systems. These results are in agreement with the hypothesis that fainter galaxies have on average higher escape fractions and account for a large part of the ionizing-photon budget (e.g., Fontanot et al. 2014). Because the bulk of the high-redshift star formation occurred in very faint galaxies (e.g., Bouwens et al. 2015), the regime of LyC emission in those galaxies needs to be studied observationally (see, e.g., Japelj et al. 2017). Deep observations of galaxies in this faint regime ( $M_{UV} \gtrsim -19$  mag) in the LyC portion of their spectra are time-demanding and therefore sparse. Fletcher et al. (2019) identified four faint LyC emission candidates and Amorín et al. (2014) studied a faint lensed galaxy resulting in a poorly constrained limit on its LyC escape fraction.

A unique way to learn more about the escape fractions of faint galaxies is to make use of long gamma-ray bursts (LGRBs). LGRBs mark the deaths of massive stars (Hjorth et al. 2003; Woosley & Heger 2006; Cano et al. 2017) and can be observed from the local to the  $z \sim 8$  Universe (Tanvir et al. 2009; Salvaterra et al. 2009). LGRBs are generally found in sub- $L^*$  galaxies at any redshift (e.g., Tanvir et al. 2012; Vergani et al. 2015; Perley et al. 2016; Palmerio et al. 2019). The progenitors are associated with young star-forming regions that contribute hugely to the ionizing photon budget inside galaxies (e.g., Ramachandran et al. 2018). A burst of gamma-ray emission is typically followed by a longer-lived, multiwavelength afterglow (e.g., Kumar & Zhang 2015). Thanks to their brightness, the spectra of the UV/optical/NIR afterglows can reveal the detailed properties of the interstellar medium (ISM) along the lines of sight toward star-forming regions, such as hydrogen column density, dust properties, metallicity and kinematics (e.g., De Pasquale et al. 2003; Thöne et al. 2014; Heintz et al. 2018; Zafar et al. 2018; Bolmer et al. 2019). In addition, the afterglow brightness allows us to search directly for possible leaking LyC in their spectra. With LGRB afterglows we can therefore study the transparency of the LGRB star-forming regions to ionizing radiation (Chen et al. 2007; Fynbo et al. 2009; Tanvir et al. 2019). Being based on afterglow spectroscopy, LGRB LyC studies do not suffer from galaxy magnitude selection and offer the advantage of a direct determination of the  $(f_{1500}/f_{900})^{\text{int}}$  ratio thanks to the intrinsically featureless (power-law) spectra of GRB afterglows. LGRB afterglows are thus a very useful tool to search for LyC leakage in very faint galaxies at  $z \sim 2 - 4$ .

In this paper we present the detection of LyC photons in the afterglow spectrum of GRB 191004B at  $z = 3.5055$ . We further investigate LyC emission in LGRB afterglow spectra, also adding to this case GRBs 050908 and 060607A ( $z = 3.3467$  and  $z = 3.0749$ , respectively), the only two previously known GRBs whose spectra show significant LyC emission (Fynbo et al. 2009; Tanvir et al. 2019). This paper is organized as follows. In Sec-

tion 2 we present the observations of GRB 191004B. We analyze them in Section 3. In Sections 4 and 5 we determine the escape fraction along the lines of sight to the three GRBs. Section 6 is focused on the Ly $\alpha$  emission modeling. All the results are discussed in Section 7, especially in the context of other known LyC-emitting galaxies. We provide our conclusions in Section 8.

All errors are reported at  $1\sigma$  confidence and all magnitudes are reported in AB unless stated otherwise. We consider a flat  $\Lambda$ CDM cosmology with the cosmological parameters provided in Planck Collaboration et al. (2016):  $H_0 = 67.8 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.308$  and  $\Omega_\Lambda = 0.692$ .

## 2. Observations

GRB 191004B was detected by the BAT instrument on the *Neil Gehrels Swift Observatory* (henceforth *Swift*) on 2019 October 4, at 21:33:41 UT (Trigger 927839; Cenko et al. 2019). The BAT light curve shows a multi-peaked shape with a duration of  $t_{90}(15 - 350 \text{ keV}) = 37.7 \pm 14.9 \text{ s}$  (Stamatikos et al. 2019). The *Swift* UV/Optical Telescope started settled observations of GRB 191004B 69 s after the trigger and refined the position to RA (J2000) =  $03^{\text{h}}16^{\text{m}}49.10^{\text{s}}$ , Dec (J2000) =  $-39^{\circ}38'03''.9$  with a 90% confidence error radius of  $\sim 0''.4$  (LaPorte & Cenko 2019). The redshift of  $z = 3.503$  was reported by D’Elia et al. (2019) based on VLT/X-shooter observations of absorption lines in the optical afterglow (see also below).

In the following we describe our ESO/VLT observations of the GRB 191004B afterglow and host galaxy obtained by the *Stargate* consortium under Program ID: 0104.D-0600 (P.I.: N. Tanvir). The log of observations is provided in Table 1.

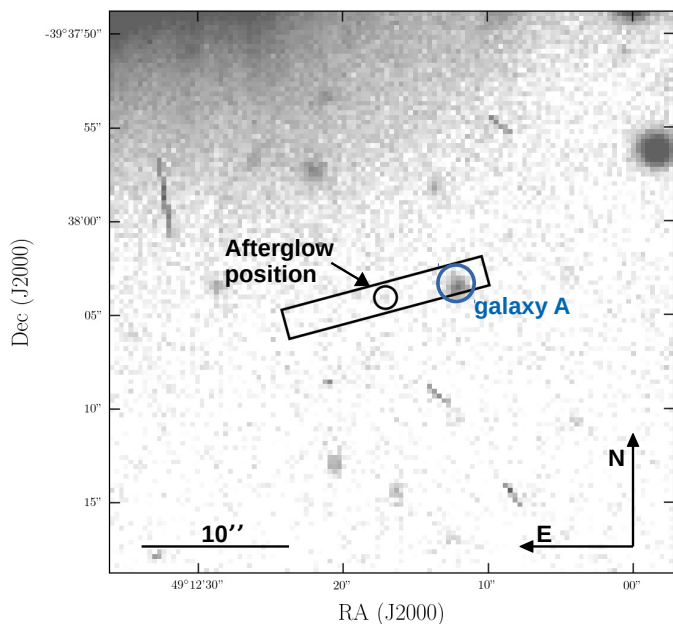
**Table 1.** Log of observations for the GRB 191004B afterglow and host galaxy.

Instrument	$\Delta t$ (day)	$T_{\text{exp}}$ (s)	Slits / Filters	Seeing (")
X-shooter	0.30	2400/2400/2400	1.0/0.9/0.9JH	0.75
X-shooter	108	5840/5440/6000	1.6/1.5/0.9	0.6
X-shooter	0.30	3x40/3x30/3x60	$g', r', z'$	0.75
FORS2	77	1800	$R\_SPECIAL$	0.7

**Notes.** The columns indicate the instrument, the time  $\Delta t$  after the GRB detection, the total exposure time  $T_{\text{exp}}$  in each slit/filter, the slit/filter configuration in the case of spectral/imaging observations and the average DIMM seeing at the time of the observation.

### 2.1. VLT/X-shooter afterglow imaging

GRB 191004B was observed with the VLT/X-shooter acquisition camera in the  $g', r', z'$  bands  $\sim 7.2$  hours after the burst. The observations were taken in a sequence of  $3 \times 40\text{s}$ ,  $3 \times 30\text{s}$  and  $3 \times 60\text{s}$  with the  $g', r'$  and  $z'$  filters, respectively. The afterglow is clearly detected in all bands at the coordinates RA(J2000) =  $03^{\text{h}}16^{\text{m}}49.14^{\text{s}}$ , Dec(J2000) =  $-39^{\circ}38'04''.09$ , with a magnitude of  $g' = 22.41 \pm 0.03$  mag,  $r' = 21.23 \pm 0.02$  mag and  $z' = 20.91 \pm 0.04$  mag. The photometry was calibrated with a single nearby star due to the small field of view. The coordinates of this star RA (J2000) =  $03^{\text{h}}16^{\text{m}}54.42^{\text{s}}$ , Dec (J2000) =  $-39^{\circ}37'41''.9$  are taken from the Gaia catalog and the photometry from the Pan-STARRS catalog.



**Fig. 1.** VLT/FORS2 image of the GRB 191004B host-galaxy field  $\sim 77.2$  days after the *Swift*/BAT trigger. The host galaxy is not detected (magnitude limit  $R \geq 26.6$  mag). The corresponding afterglow position (error radius  $r_e = 0''.6$ , as observed in the X-shooter  $r'$  image  $\sim 7.2$  hours after the burst) is indicated by a black circle. The slit at PA =  $-75^\circ$  used for the X-shooter afterglow spectroscopy is also overplotted. Galaxy A corresponds to the intervening system at  $z = 2.137$  identified in the VLT/X-shooter spectrum.

## 2.2. VLT/X-shooter afterglow spectroscopy

The first epoch spectrum of GRB 191004B was obtained with the VLT/X-shooter echelle spectrograph (Vernet et al. 2011) just after the  $g'$ ,  $r'$ ,  $z'$  imaging,  $\sim 7.2$  hours after the *Swift*/BAT trigger. The observation consisted of a total integration time of 2400 s in each arm using the  $1''.0/0''.9/0''.9$  JH slits at a position angle of PA =  $-75^\circ$  (see Fig. 1). The observation was carried out under good conditions with an average seeing of  $\sim 0''.75$  and an airmass of 1.1. We used a nodding scheme with a  $5''.0$  nodding throw. This observation covers the afterglow emission of the GRB across the wavelength range 300 - 2100 nm.

We reduced the X-shooter data using version 3.3.5 of the X-shooter data reduction pipeline (Modigliani et al. 2010). All observations of the afterglow and host galaxy, the telluric stars, and the spectrophotometric standards were reduced in the same way. Before processing the spectra through the pipeline, the cosmic-ray hits and bad pixels were removed following the method of van Dokkum (2001). Then, we subtracted the bias from all raw frames and divided them by the master flat field. We traced the echelle orders and calibrated the data in spatial and wavelength units using arc-line lamps. The flux calibration was done using spectrophotometric standards (Vernet et al. 2009) and a correction for mirror flexure was applied. Lastly, the sky-subtraction and the rectification and merging of the orders was done to obtain the final two-dimensional spectra. To optimally select the extraction region we chose the spatial extension of the afterglow continuum independently in each of the three (UVB, VIS, NIR) arms of the spectrograph. We corrected the 1D spectra for Galactic extinction using the Milky-Way extinction curve of Pei (1992) (assuming the ratio of total-to-selective extinction  $R_V = 3.08$ ) and the extinction map of Schlafly & Finkbeiner (2011). This final data set typically has a signal to noise ratio  $S/N > 3$  per

spectral bin and a spectral resolution of  $R \sim 6700/11000/7800$  in the UVB/VIS/NIR arms, respectively. The resolution was estimated from the average seeing reached during the observation corrected for airmass and wavelength dependence and by interpolating the nominal values available on the ESO instrument website<sup>1</sup>. We also compare these values to the resolutions obtained using the relation derived in Selsing et al. (2019) and we find a good agreement between both. We clearly detect the afterglow continuum from 3675 Å to the end of the X-shooter spectrum. We identify the Lyman-alpha ( $\text{Ly}\alpha$ ) emission and absorption line at  $\sim 5474$  Å and detect the metal absorption lines described in detail in Sect. 3.1.

## 2.3. Host-galaxy imaging and spectroscopy

We observed the field of GRB 191004B with the VLT/FORS2 (Appenzeller et al. 1998) camera in imaging mode on 2019 December 21, at 03:25:03 UT ( $\sim 77$  days after the *Swift*/BAT trigger). In the 30 min  $R$ -band image (see Fig. 1), we do not detect the host galaxy. The magnitude limit is  $R \geq 26.6$  mag.

A further VLT/X-shooter spectrum was taken at the afterglow position, once the afterglow faded ( $\sim 108$  days after the *Swift*/BAT trigger), using the blind offset technique to look for host galaxy emission lines. The average seeing and airmass during the observation were  $\sim 0''.6$  and 1.2, respectively. The observations were divided in two OBs for a total integration time of 5840/5440/6000 s in the UVB/VIS/NIR arm, respectively. We used the  $1''.6/1''.5/0''.9$  slits at a position angle of  $-40^\circ$  and a nodding scheme with a  $5''.0$  nodding throw. The data reduction was carried out in a similar fashion as described in Sect. 2.1. We corrected the 1D spectra for Galactic extinction as for the afterglow spectrum (Sect. 2.2). We clearly detect the  $\text{Ly}\alpha$  emission line at the same wavelength as in the afterglow spectrum but we do not detect other nebular emission lines. Unlike the afterglow spectra, no continuum is detected in the host-galaxy spectra, so to optimally select the extraction region we chose the spatial extension of the  $\text{Ly}\alpha$  emission line and applied this 1D extraction throughout the whole spectrum. We provide flux and upper limits of the lines in Sect. 3.1. The good seeing obtained during this observation allows us to reach resolutions of  $R \sim 7800/12900/9200$  in the UVB/VIS/NIR arms, respectively, calculated similarly as in Sect. 2.2.

## 3. Data analysis

### 3.1. H I and metal absorption lines

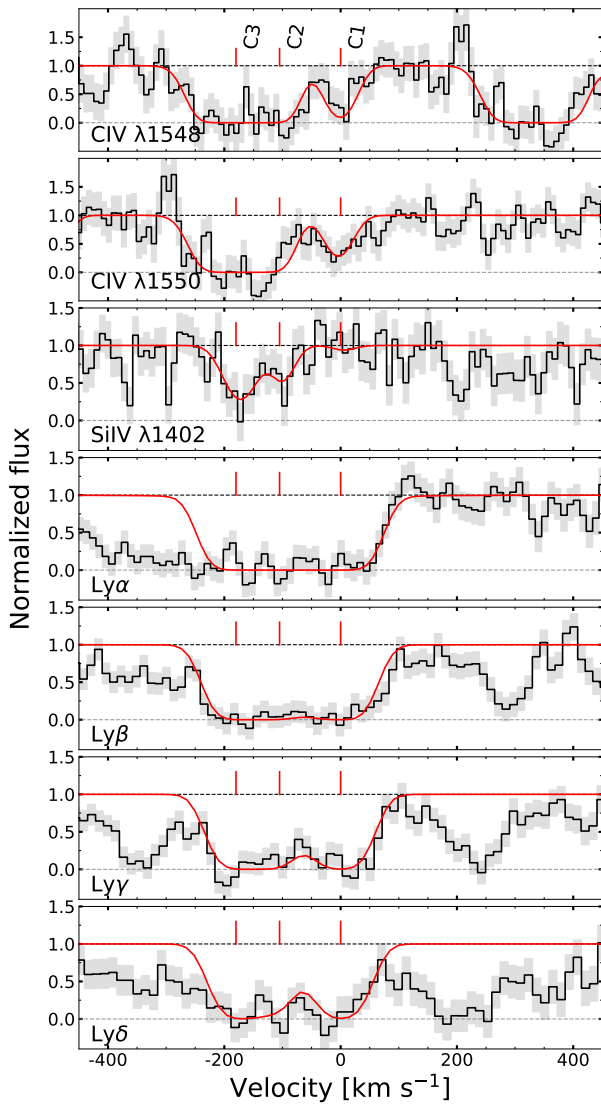
The X-shooter afterglow spectrum reveals  $\text{Ly}\alpha$ ,  $\text{Ly}\beta$ ,  $\text{Ly}\gamma$ , and  $\text{Ly}\delta$ , Si IV and C IV absorption lines from the host galaxy gas (D'Elia et al. 2019) (see Fig. 2).

Firstly, we fit the saturated Si IV and C IV absorption with three components using VoigtFit (Krogager 2018). We determine lower limits on the C IV and Si IV column densities of  $\log(N_X/\text{cm}^{-2}) > 15.6$  and 14.1, respectively, and a redshift of  $z = 3.5028, 3.5039$ , and  $3.5055$  for the three components (corresponding to relative velocity from the reddest component of  $\sim -105$  and  $-180$  km s $^{-1}$ , respectively). The Doppler parameters are  $b = 25, 20, 30$  km s $^{-1}$  (from the red component to the blue one, respectively).

The high-ionization metal lines such as C IV and Si IV do not necessarily originate from the same place as the neutral gas (e.g.,

<sup>1</sup> <https://www.eso.org/sci/facilities/paranal/instruments/xshooter/inst.html>





**Fig. 2.** C IV, Si IV and Lyman series absorptions of the GRB 191004B host galaxy identified in the VLT/X-shooter afterglow spectrum. The gray regions correspond to the uncertainty on the flux and the red line to the best-fit model, obtained using three components. Zero corresponds to the reddest component at  $z = 3.5055$  named C1. C2 and C3 are the two other components detected at  $-105$  and  $-180$   $\text{km s}^{-1}$ , respectively.

Heintz et al. 2018). However, since they are the only detected for the GRB system, we fixed these components and their redshift to fit the Lyman lines. The Lyman lines are saturated, therefore their fitting could not give a precise estimate of the hydrogen column density. Fixing the  $b$  parameters to those inferred from the corresponding metal line components, we find an acceptable fit for  $\log(N_{\text{H I}}/\text{cm}^{-2}) = 17.2 \pm 0.3$ . A fit letting the redshift of the components and the  $b$  parameters free to vary would give consistent  $N_{\text{H I}}$  values. This is one of the lowest hydrogen column densities measured among the LGRB afterglow sample (Tanvir et al. 2019) and places the absorbing system on the boundary between being a Ly $\alpha$  forest absorber ( $\log(N_{\text{H I}}/\text{cm}^{-2}) < 17$ ; Rauch 1998) or a Lyman-limit system (LLS,  $17 < \log(N_{\text{H I}}/\text{cm}^{-2}) < 20.3$ ; Péroux et al. 2003). The reddest component is the richest in neutral hydrogen, whereas the bluest component is the strongest for the high-ionization lines.

We do not detect the other common singly ionized metal lines such as Fe II, Mg II, Al II or Si II (Christensen et al. 2011; de

Ugarte Postigo et al. 2012), that are typically detected in LGRB afterglow spectra. The part of the spectrum at the corresponding wavelengths of most of the low-ionization lines is quite noisy, and/or contaminated by residual sky lines. We place approximate upper limits on their rest-frame equivalent widths of  $\sim 0.4$  Å. We can determine a  $3\sigma$  upper limit on the C II column density of  $\log(N_{\text{H I}}/\text{cm}^{-2}) > 14.7$ . Low column densities of low-ionization metal transitions are also found for a part of the LLS sample detected through quasar spectroscopy (see, e.g., Fox et al. 2013; Cooper et al. 2015), as well as in the other few LGRBs having  $\log(N_{\text{H I}}/\text{cm}^{-2}) < 19$  (see, e.g., Thöne et al. 2011). In these systems a significant fraction of the gas is ionized, whereas usually LGRB hosts show higher H I absorption values (median of  $\log(N_{\text{H I}}/\text{cm}^{-2}) = 21.6$ ; Tanvir et al. 2019), typical of systems probing neutral gas. At such low H I column densities, even for approximately solar metallicity gas, the strongest low-ionization metal absorption features can have column densities below the detection threshold. De Cia et al. (2011) found very low values for the column densities of low-ionization metals detected in GRB 070125 afterglow spectra, at  $z = 1.5477$ . They conclude that the LGRB was likely located in the outskirts of a massive star-forming region inside a faint and small host galaxy (see also Utdike et al. 2008; Cenko et al. 2008).

We detect Fe II, Mg II absorption lines and [O III] emission corresponding to an intervening absorber at  $z = 2.137$ . This system is at  $3''.8$  from the afterglow position and corresponds to galaxy A in Fig. 1. Its emission trace in the 2D spectrum covers  $1''.5$ . We do not detect the presence of another galaxy brighter than  $R=26.6$  mag in a radius of  $3''.0$  around the position of the GRB 191004B afterglow, corresponding to 22 kpc at this redshift. Therefore, we can rule out contamination of the observed LyC emission from foreground interlopers.

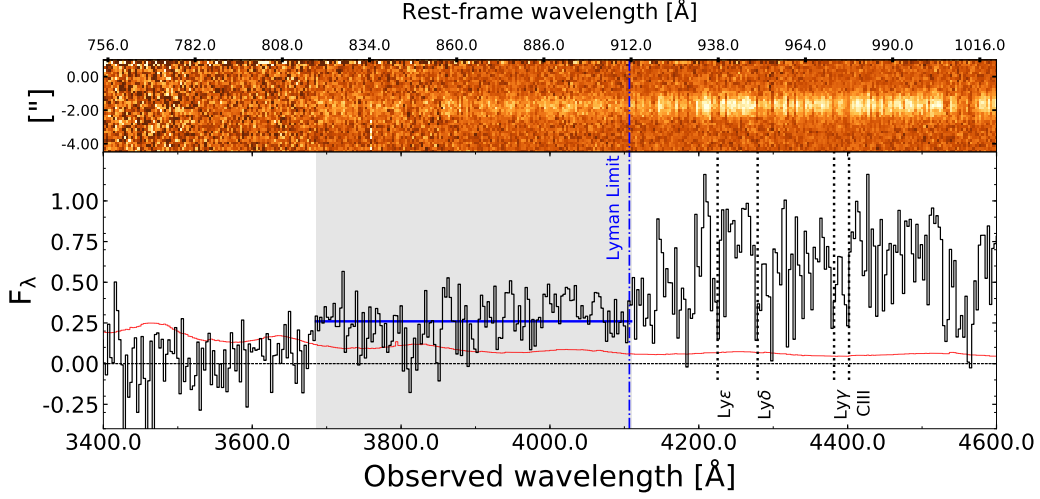
### 3.2. Lyman-continuum leakage

The afterglow spectrum of GRB 191004B clearly presents residual flux in the observed wavelength range  $3675.00$ – $4106.73$  Å ( $816$ – $912$  Å rest-frame), see Fig. 3. We measure an average flux density of  $f_{900}^{\text{obs}} = (2.6 \pm 0.2) \times 10^{-18}$  erg  $\text{s}^{-1} \text{cm}^{-2} \text{Å}^{-1}$  over this wavelength range.

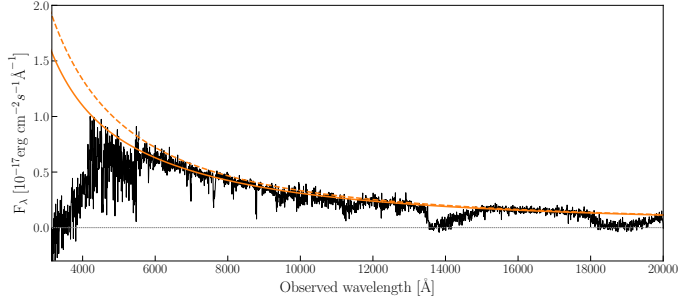
We detect several Ly $\alpha$  absorbers in the Ly $\alpha$  forest from  $z = 2.8$  to  $z = 3.45$ , with hydrogen column densities spanning  $\log(N_{\text{H I}}/\text{cm}^{-2}) \sim 15$ – $17$ . The lack of LyC emission below  $\sim 3650$  Å is due to the Lyman limit absorption by some of these systems, and the dimming of the LyC emission around  $3800$  Å corresponds to the Lyman limit of a system at  $z = 3.19$  and to the likely strong Ly $\alpha$  absorption of Galaxy A (see Sect. 3.1).

The non-ionizing UV flux density measured in the rest-frame range  $1480$ – $1520$  Å is  $f_{1500}^{\text{obs}} = (5.5 \pm 0.1) \times 10^{-18}$  erg  $\text{s}^{-1} \text{cm}^{-2} \text{Å}^{-1}$ . Together they correspond to a non-ionizing to ionizing flux ratio of  $(f_{1500}/f_{900})^{\text{obs}} = 2.1 \pm 0.2$ . This ratio will be used to measure the escape fraction of ionizing radiation in Sect. 4.1.

The host-galaxy spectrum does not show LyC emission above a  $3\sigma$  flux density limit of  $4.5 \times 10^{-18}$  erg  $\text{s}^{-1} \text{cm}^{-2} \text{Å}^{-1}$ . A rough binning of the spectrum in the  $820$ – $910$  Å range allows to place a more stringent  $3\sigma$  upper limit on its LyC flux density to  $6.0 \times 10^{-19}$  erg  $\text{s}^{-1} \text{cm}^{-2} \text{Å}^{-1}$ .



**Fig. 3.** Part of the afterglow UVB spectrum of the GRB 191004B covering the 912 Å region (rest-frame). Top: 2D spectrum showing the detection of residual emission below the Lyman limit. Bottom: 1D spectrum (black line) with flux density  $F_\lambda$  in units of  $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ Å}^{-1}$ . Its associated error spectrum ( $1\sigma$ ) is in red. The grey shaded region marks the rest-frame range 816 - 910 Å where residual flux below 912 Å is detected. The solid blue line represents the corresponding average LyC flux density of  $f_{900} = (2.6 \pm 0.2) \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ Å}^{-1}$  over this range. The spectrum is rebinned at 3 Å per pixel.



**Fig. 4.** Fit of the dust extinction to the afterglow spectrum of GRB 191004B (black line). The solid orange line corresponds to the best-fit extinction (solid orange line) with a spectral index  $\beta_o = 0.49 \pm 0.03$  and an SMC extinction law with  $A_V = 0.03^{+0.02}_{-0.02}$  mag. See Sect. 3.3 for more details. The orange dashed line is for the intrinsic spectrum corrected for extinction.

### 3.3. Line-of-sight extinction

We measured the host-galaxy extinction along the line of sight to the GRB directly from the afterglow spectrum corrected for Galactic extinction (see Sect. 2.2) following the procedure as described in Japelj et al. (2015) and Zafar et al. (2018). Due to the sparse X-ray afterglow observations around the epoch of the X-shooter spectrum, we did not include X-ray data into this fit. We ignored the spectral region blue-wards of the Ly $\alpha$  emission/absorption feature in the fit, since it is significantly affected by Ly $\alpha$  forest absorption. The edges of the VIS and NIR spectra and the regions affected by telluric absorption were also masked out. We cleaned the spectra of absorption lines by recursively fitting a polynomial to the continuum and removing everything that deviate from the fit by more than  $3.5\sigma$ . Following the procedure of Japelj et al. (2015), we fitted the spectrum with a power-law function modulated by dust extinction assuming the Milky Way and Magellanic Clouds extinction curves from Pei (1992). We found that the spectrum can be well described by a power-law with small extinction. The best fit gives an optical spectral index  $\beta_o = 0.49 \pm 0.03$ , an extinction  $A_V = 0.03 \pm 0.02$  mag

and  $\chi^2/\text{dof} = 1.007$  for the Small Magellanic Cloud extinction curve. We present the best fit (solid orange line) together with the corresponding intrinsic power-law corrected for extinction (dashed line) in Fig. 4.

### 3.4. Ly $\alpha$ emission

We detect the Ly $\alpha$  emission line in both the X-shooter afterglow and host-galaxy spectra (see Fig. 5). The emission line is asymmetric. We fit its profile with a skewed Gaussian parametrized as in Claeysens et al. (2019); Matthee et al. (2019),

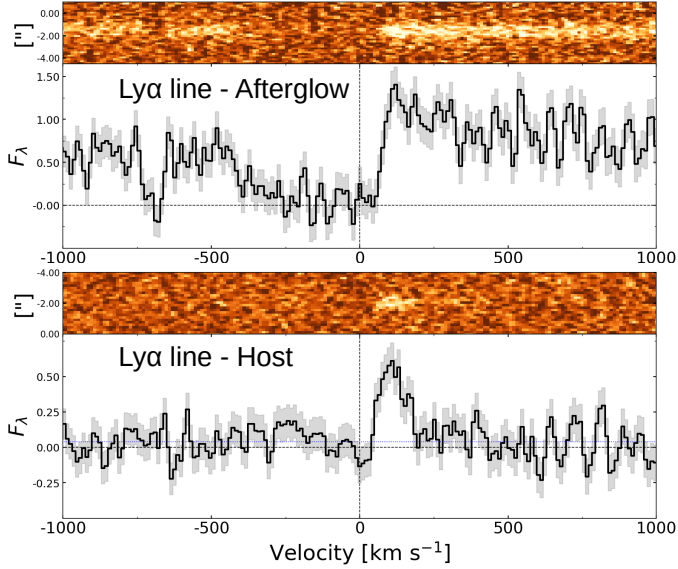
$$f(\lambda) = A \exp \left[ -\frac{(\lambda - \lambda_0)^2}{2(a_{\text{asym}}(\lambda - \lambda_0) + d)^2} \right], \quad (1)$$

where  $A$  is the amplitude of the Ly $\alpha$  line;  $\lambda_0$  is its peak wavelength;  $a_{\text{asym}}$  is the measure of asymmetry of the line and  $d$  the parameter that controls the line width. These parameters are left free during the fitting with broad uniform priors. In order to determine both a robust fit and its associated uncertainty, we use the Markov Chain Monte Carlo sampler emcee (Foreman-Mackey et al. 2013) to maximize the likelihood function using Eq. (1). The best fit of the X-shooter Ly $\alpha$  line detected in the host spectrum provides a peak position of  $\lambda_0 = 5478.85^{+0.13}_{-0.12}$  Å, an asymmetry of  $a_{\text{asym}} = 0.33^{+0.05}_{-0.07}$  and a width parameter of  $d = 0.6985^{+0.110}_{-0.103}$ . We convert the latter into FWHM according to the analytic expression:

$$FWHM = \frac{2 \sqrt{2 \log(2)} d}{1 - 2 \log(2) a_{\text{asym}}^2}. \quad (2)$$

The corresponding rest-frame FWHM of the line is  $FWHM_0 = 104 \pm 30 \text{ km s}^{-1}$ . The flux of the Ly $\alpha$  line corrected for Galactic extinction is  $F(\text{Ly}\alpha) = (1.0 \pm 0.1) \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ . This corresponds to a Ly $\alpha$  luminosity of  $L_{\text{Ly}\alpha} = (1.18 \pm 0.12) \times 10^{42} \text{ erg s}^{-1}$  and a rest-frame equivalent width of  $EW_0(\text{Ly}\alpha) = 7.4 \pm 2.6 \text{ Å}$ .

We convert this Ly $\alpha$  flux to a star formation rate of  $\text{SFR}(\text{Ly}\alpha) \approx 1 M_\odot \text{ yr}^{-1}$  by using the recombination factor 8.7 from the case



**Fig. 5.** Part of the 2D and 1D UVB spectra of the GRB 191004B in velocity space, centered on the Ly $\alpha$  line at  $z = 3.5055$ . The flux density  $F_\lambda$  is in units of  $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ . Top: VLT/X-shooter spectrum of the afterglow  $\sim 7.2$  hours after the burst detection. The Ly $\alpha$  line is visible in the red wing of the absorption line. Bottom: VLT/X-shooter spectrum of the host galaxy. The Ly $\alpha$  emission line is clearly detected at  $113 \pm 30 \text{ km s}^{-1}$ .

B of the theory of recombination (Brocklehurst 1971) and the relation from Kennicutt (1998), scaled to the Chabrier (2003) initial mass function. Nevertheless, we know that this SFR can be largely underestimated due to the complex scattering and destruction process of Ly $\alpha$  photons in neutral gas and dust. Therefore the SFR derived here only represents a lower limit.

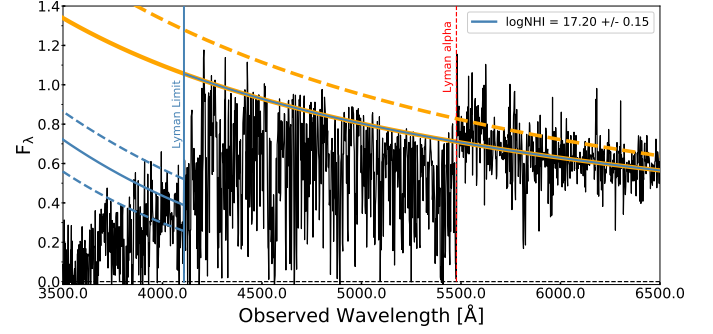
### 3.5. Other emission lines

We do not detect other emission lines at the afterglow position in the afterglow and host-galaxy spectra. We determine  $3\sigma$  upper limits of  $0.3 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$  for H $\beta$  and [OII] $_{\lambda 3727}$ , and  $1.0 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$  for [OIII] $_{\lambda 5007}$ . By assuming that the negligible dust extinction in the line of sight is also valid for the integrated host-galaxy dust content, the H $\beta$  upper limit can be converted to an SFR  $< 4.7 \text{ M}_\odot \text{ yr}^{-1}$  considering the intrinsic Balmer decrement  $H\alpha/H\beta = 2.86$  (Osterbrock 1989) and using the relation from Kennicutt (1998), scaled to the Chabrier (2003) initial mass function (IMF). This value is consistent with the lower limit derived from the Ly $\alpha$  line flux and allows us to put an additional constraint on the SFR. This implies an SFR of a few  $\text{M}_\odot \text{ yr}^{-1}$  and an escape fraction of Ly $\alpha$  photons  $f_{\text{esc}}(\text{Ly}\alpha) > 0.13$ .

## 4. LyC escape fraction

The LyC escape fraction is the ratio of the observed flux below the Lyman limit, corrected for the Milky Way extinction and the IGM transmission, and the intrinsic flux. Usually in galaxy studies this last quantity is inferred from the galaxy SED (and therefore it is model dependent), and the IGM transmission used is an average value obtained from line-of-sight simulations.

Taking advantage of afterglow spectroscopy it is possible to determine  $f_{\text{esc}}$  (i): directly using the  $N_{\text{HI}}$  value derived by fitting the Ly $\alpha$  absorption line or the envelope of the residual



**Fig. 6.** Afterglow spectrum of the GRB 191004B covering the LL and Ly $\alpha$  line. The flux of the escaping LyC photons obtained for  $\log(N_{\text{HI}}/\text{cm}^{-2}) = 17.20 \pm 0.15$  corresponds to the blue solid and dashed lines below the LL. The orange solid/dashed line corresponds to the fit of the afterglow continuum with/without correction for dust extinction. The flux density  $F_\lambda$  is in units of  $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ . The spectrum is rebinned to  $2 \text{ \AA}$  per pixel.

flux below the Lyman limit (see Sect. 4.1); or (ii): through the  $(f_{1500}/f_{900})^{\text{int}}$  ratio and IGM simulation (see Sect. 4.2), as commonly done in galaxy studies. The former case is a *direct* determination from the data. The fact that the afterglow continuum is a simple power-law is an advantage (in both cases) because the intrinsic  $f_{900}$  is much less model dependent than in galaxy studies.

### 4.1. Direct determination

Following Siana et al. (2007), we define  $f_{\text{esc}}$  as the fraction of emitted Lyman continuum photons that escapes into the IGM. This corresponds to the formula,

$$f_{\text{esc}} = \exp[-\tau_{\text{LL}}(\lambda)], \quad (3)$$

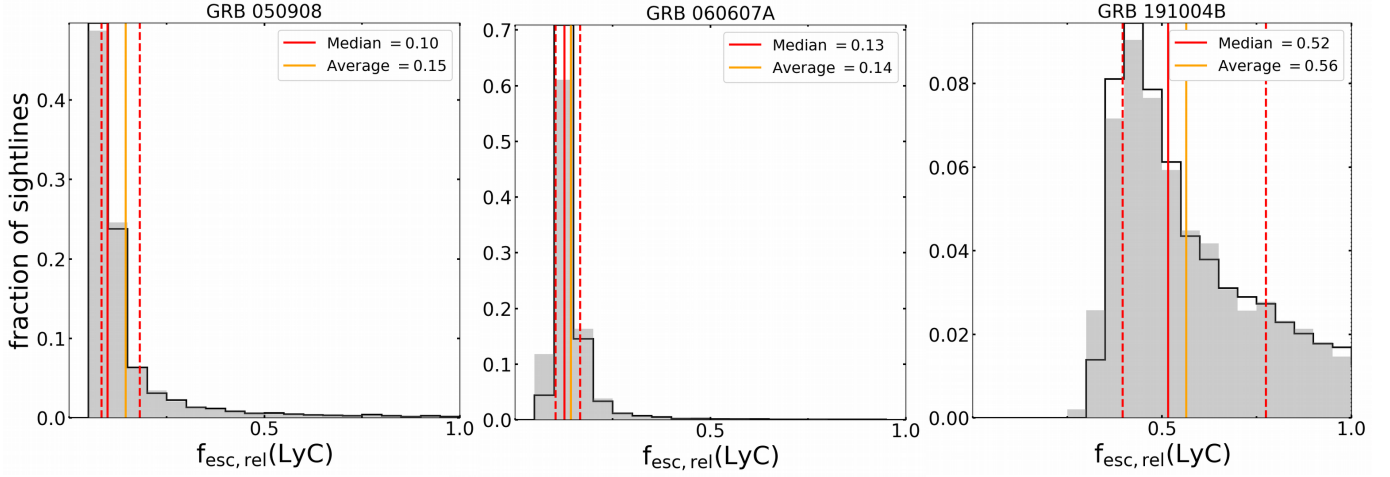
where  $\tau_{\text{LL}}$  is the optical depth at wavelengths below the Lyman limit and is given by (see also Prochaska et al. 2010),

$$\tau_{\text{LL}}(\lambda) \approx \frac{N_{\text{HI}}}{10^{17.2} \text{ cm}^{-2}} \left( \frac{\lambda}{\lambda_{\text{LL}}} \right)^3, \quad (4)$$

where  $\lambda_{\text{LL}} \approx 912 \text{ \AA}$  is the Lyman limit wavelength. In principle we can directly use the value of  $N_{\text{HI}}$  obtained by the fit of the HI absorption. In the case of GRB 191004B, the HI column density belongs to the flat part of the HI curve of growth, therefore it cannot be determined precisely. Considering the value and errors determined in Sect. 3.1, we find  $f_{\text{esc}} = 0.43^{+0.22}_{-0.24}$ . The large error range reflects the  $N_{\text{HI}}$  uncertainties.

We can determine more precise  $N_{\text{HI}}$  and  $f_{\text{esc}}$  values in the following way. The observed afterglow spectrum bluewards of the Lyman limit is the afterglow continuum (a power-law modulated by extinction, i.e. the solid line in Fig. 4), heavily absorbed by the neutral hydrogen in the host galaxy (Eq. (4)), and furthermore absorbed to a much lesser extent by Lyman alpha forest at the corresponding redshift. From the shape of the observed spectrum and the modelled afterglow continuum we can use Eq. (4) to directly determine the opacity in the host galaxy's line-of-sight by varying the  $N_{\text{HI}}$  (see Fig. 6). Because the observed spectrum is partly absorbed by the Lyman alpha forest we don't aim to match the spectral continuum but instead its envelope. The best match is obtained with  $\log(N_{\text{HI}}/\text{cm}^{-2}) = 17.20 \pm 0.15$ , which corresponds to  $\langle f_{\text{esc}} \rangle = 0.43^{+0.12}_{-0.13}$ . The errors reflect the uncertainty on the absorption due to the forest. If we take into account the





**Fig. 7.** Distribution of *relative* escape fraction ( $f_{\text{esc,rel}}(\text{LyC})$ ) for GRB 050908, GRB 060607A and GRB 191004B (from left to right) with and without uncertainty on the observed measure of  $(f_{1500}/f_{900})^{\text{obs}}$  from the afterglow spectrum (respectively gray and black distribution). The high values with  $f_{\text{esc,rel}}(\text{LyC}) > 1$  are non-physical and due to low IGM transmission, hence we restrict the statistics and axis range to 0-1 for  $f_{\text{esc,rel}}(\text{LyC})$ . For GRB 191004B, 57% of the gray distribution is at  $f_{\text{esc,rel}}(\text{LyC}) \leq 1$  while the fraction is 92% and 96% for GRB 050908 and GRB 060607A, respectively. The solid red and orange lines in each panel represent the median and average value of this  $f_{\text{esc,rel}}(\text{LyC}) \leq 1$  part of the distribution. The dashed red lines represent the corresponding 16th and 84th percentiles. The  $f_{\text{esc,abs}}$  are similar to the values quoted here for GRB 050908 due to the null extinction hypothesis. For GRB 191004B and GRB 060607A the corresponding absolute values are derived in Sect. 4.2 and 5.

effect of dust extinction on the afterglow continuum, we obtain  $\langle f_{\text{esc,abs}} \rangle = 0.35^{+0.10}_{-0.11}$ .

We can also determine a lower limit to  $f_{\text{esc}}$  and  $N_{\text{HI}}$  from the observed continuum over the same wavelength range, as it corresponds to the flux of the escaping photons dimmed by the IGM opacity. We obtain  $\log(N_{\text{HI}}/\text{cm}^{-2}) < 17.4$ ,  $f_{\text{esc}} > 0.25$ , and  $f_{\text{esc,abs}} > 0.22$  taking dust extinction into account.

#### 4.2. Constraints from IGM simulations

In galaxy studies, the common way to estimate the fraction of escaping ionizing photons is by calculating the ratio between the observed fraction of escaping LyC photons (900 Å), corrected for IGM transmission, relative to the fraction of escaping non-ionizing photons (1500 Å) (Steidel et al. 2001; Siana et al. 2007). The *relative* escape fraction can be expressed as:

$$f_{\text{esc,rel}} = \frac{(f_{1500}/f_{900})^{\text{int}}}{(f_{1500}/f_{900})^{\text{obs}} T_{900}^{\text{IGM}}}, \quad (5)$$

where  $(f_{1500}/f_{900})^{\text{int}}$  is the intrinsic flux density ratio,  $(f_{1500}/f_{900})^{\text{obs}}$  is the observed flux density ratio and  $T_{900}^{\text{IGM}} = e^{-\tau_{\text{IGM}}}$  is the IGM transmission factor at 900 Å along the sightline, and  $\tau_{\text{IGM}}$  the line-of-sight opacity to the IGM for LyC photons. As shown in Siana et al. (2007), this formula is equivalent to Eq. (3).

The *absolute* escape fraction,  $f_{\text{esc,abs}}$  is the ratio of the escaping to intrinsic ionizing flux density. Knowing the dust attenuation  $A_{1500}$ ,  $f_{\text{esc,abs}}$  can be written as (e.g., Inoue et al. 2005; Siana et al. 2007),

$$f_{\text{esc,abs}} = f_{\text{esc,rel}} 10^{-0.4 A_{1500}}. \quad (6)$$

The  $(f_{1500}/f_{900})^{\text{int}}$  ratio is very difficult to constrain observationally and is usually estimated by using spectral synthesis models (e.g., Bruzual & Charlot 2003). Therefore, this value has generally significant uncertainties due to the assumptions on which these models rely, stellar population age, metallicity, star-formation history and IMF. In our case, we have the possibility

to directly measure the intrinsic flux density ratio as the intrinsic afterglow emission corresponds to a single power-law. This power-law is estimated as described in Sect. 3.3 and extrapolated to the ionizing UV domain. We then calculate the average intrinsic flux density in the same rest-frame range as for the observed flux densities ( $f_{1500}^{\text{obs}}$  and  $f_{900}^{\text{obs}}$ , Sect. 3.2). We finally derive an intrinsic flux ratio of  $(f_{1500}/f_{900})^{\text{int}} = 0.43$  from the best fit corrected for extinction (dashed orange curve in Figure 4).

To recover the  $T_{900}^{\text{IGM}}$  factor, we simulated a large number of sightlines using observational constraints on the properties of the IGM probed in intergalactic absorbers. We used the same Monte Carlo (MC) simulation as in Japelj et al. (2017) and described in Vanzella et al. (2015) to generate 10 000 line-of-sight transmissions from a source at  $z = 3.5$ . For each simulated sightline we calculate the average transmission over the same rest-frame range as used for  $f_{1500}^{\text{obs}}$  and  $f_{900}^{\text{obs}}$ , we then correct the observed ratio  $(f_{1500}/f_{900})^{\text{obs}}$  for this quantity and calculate  $f_{\text{esc,rel}}(\text{LyC})$ . In order to take into account the uncertainty on  $(f_{1500}/f_{900})^{\text{obs}}$ , for each of the 10 000 sightlines we randomly select a value in a normal distribution defined by its measurement (center) and its error (width). The resulting distribution of  $f_{\text{esc,rel}}$  is shown in grey in Fig. 7. In the case of GRB 191004B, the best fit of the afterglow provides an estimate of the extinction and hence the possibility to calculate the absolute escape fraction. In the following, we report the absolute values derived using Eq. (6) but for a comparison we also provide the relative ones, which do not rely on a dust-extinction model (see Sect. 3.3).

For the estimation of the escape fraction we proceed similarly to what is done in Shapley et al. (2016) in their study of the direct detection of LyC emission from a galaxy at  $z \sim 3$ . We only keep the lines of sight resulting in an escape fraction of  $f_{\text{esc}} \leq 1$ , since the values higher than one are not physical. We find that 57% of the initial 10 000 sightlines are at  $f_{\text{esc}} \leq 1$  and we measure that 95% of the distribution is at  $f_{\text{esc,abs}} \geq 0.31$  ( $f_{\text{esc,rel}} \geq 0.35$ ). The median escape fraction is  $f_{\text{esc,abs}} = 0.46^{+0.23}_{-0.11}$  ( $f_{\text{esc,rel}} = 0.52^{+0.26}_{-0.12}$ ), with an average value of  $\langle f_{\text{esc,abs}} \rangle = 0.50$  ( $\langle f_{\text{esc,rel}} \rangle = 0.56$ ), see also Table 2.



**Table 2.** Escape fraction of ionizing photons along the line of sight for the three GRBs.

GRB	$z_{\text{syst}}$	$f_{\text{esc,abs}}^{\text{HI}} (\text{rel})$	$f_{\text{esc,abs}} (\text{rel})$	$f_{\text{esc,abs}}^{\text{lim}}$
050908	3.3467	$0.09^{+0.06}_{-0.04} (*)$	$0.10^{+0.09}_{-0.02} (*)$	$>0.07$
060607A	3.0749	$0.20^{+0.05}_{-0.05} (0.45^{+0.13}_{-0.13})$	$0.08^{+0.03}_{-0.02} (0.13^{+0.04}_{-0.03})$	$>0.05$
191004B	3.5055	$0.35^{+0.10}_{-0.11} (0.43^{+0.12}_{-0.13})$	$0.46^{+0.23}_{-0.11} (0.52^{+0.26}_{-0.12})$	$>0.31$

**Notes.** The columns correspond to the name of the GRB;  $z_{\text{syst}}$ : the systemic redshift of the host galaxy;  $f_{\text{esc,abs}}^{\text{HI}} (\text{rel})$ : the absolute (abs) and relative (rel) escape fraction determined from the estimation of the HI column density, see Sect. 4.1;  $f_{\text{esc,abs}} (\text{rel})$ : the absolute (abs) and relative (rel) lower limit of the escape fraction of ionizing photons. The relative value is given between brackets when a non-zero extinction is considered, otherwise (\*) means same value as the absolute one (see Sect. 4.2);  $f_{\text{esc,abs}}^{\text{lim}}$ : the absolute lower limit of the escape fraction of ionizing photons (see Sect. 4.2).

## 5. LyC leakage among LGRB host galaxies

To date, in addition to GRB 191004B presented in this paper, there are only two more LGRBs for which direct LyC emission was detected, namely GRB 050908 and GRB 060607A. Their afterglow spectra have already been presented in Fynbo et al. (2009) and briefly discussed in Tanvir et al. (2019). Here, we summarize the main known features of these LGRB afterglows and host galaxies and derive their  $f_{\text{esc}}(\text{LyC})$ , following the procedures of Sect. 4. We stress that, apart from their afterglow optical spectrum showing low HI column densities associated with the LGRB host galaxy, the general properties of these LGRBs (prompt emission, energetic, afterglow light curves) are similar to that of the general LGRB population (see, e.g., Kann et al. 2010).

GRB 050908 occurred at  $z = 3.3467$ . Its afterglow spectrum was observed with VLT/FORS1 (Fynbo et al. 2009), KeckII/Deimos and Gemini North/GMOS (Chen et al. 2007), and showed very strong high- and low-ionization absorption lines at the GRB redshift. In this paper we use the spectrum obtained by VLT/FORS. In addition to the absorptions associated with the GRB host galaxy, two strong Mg II intervening absorbers are identified in the GRB afterglow spectrum at  $z = 0.6915$  and  $1.4288$ . The emission counterparts of these absorbers have been identified by Schulze et al. (2012) a few tens of kpc away from the GRB host. Deep *HST-F775W* images of the field allowed the detection of the GRB host galaxy with  $m_{AB} = 27.67$  mag (Hjorth et al. 2012; Blanchard et al. 2016). The VLT/FORS spectrum covers the ionizing domain at wavelengths  $> 818 \text{ \AA}$  rest-frame. A sufficient signal-to-noise ratio is only reached down to  $880 \text{ \AA}$ . We measure the ionizing ( $f_{900}^{\text{obs}}$ ) and non-ionizing UV ( $f_{1500}^{\text{obs}}$ ) flux density from this spectrum and find a  $(f_{1500}/f_{900})^{\text{obs}}$  ratio of  $6.9 \pm 0.5$ . Considering a null extinction ( $A_V = 0$  mag), according to the results published in Vergani (2008), we fit the afterglow spectrum of GRB 050908 with a simple power-law and following the method described in Sect. 4.1, we derive  $(f_{1500}/f_{900})^{\text{int}} = 0.35$  with  $f_{900}^{\text{int}}$  measured over  $880\text{--}910 \text{ \AA}$ . Similarly to GRB 191004B we can derive the absolute escape fraction of ionizing photons which is equal to the relative one in this case, since no extinction correction is applied. The distribution of  $f_{\text{esc,rel}}(\text{LyC})$  is presented in Fig. 7 and shows that 95% of the distribution is at  $f_{\text{esc,abs}} \geq 0.07$  and 92% with  $f_{\text{esc,abs}} \leq 1$ . The median escape fraction for the latter 92% is

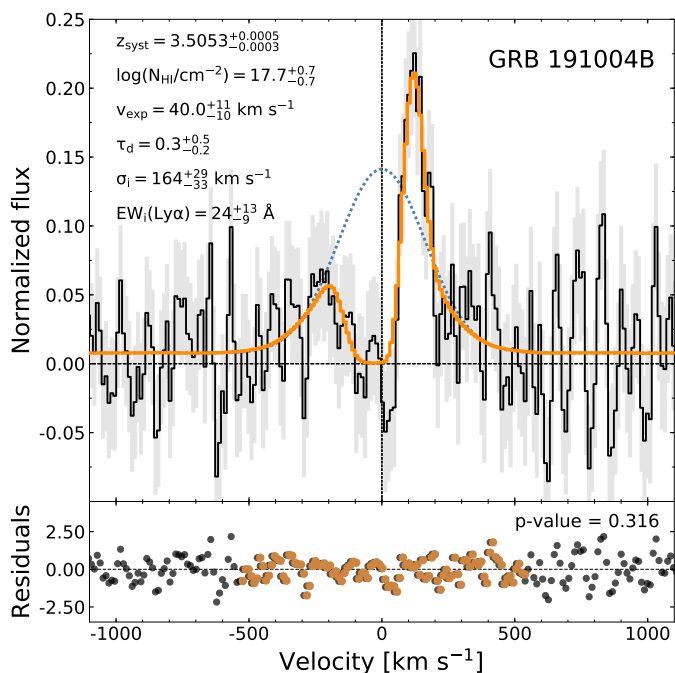
$0.10^{+0.08}_{-0.02}$ , with an average value of  $\langle f_{\text{esc,abs}} \rangle = 0.15$ . This estimation is fully consistent with the value of  $f_{\text{esc}} = 0.08^{+0.05}_{-0.04}$  derived by Tanvir et al. (2019) from the HI column density ( $\log(N_{\text{HI}}/\text{cm}^{-2}) = 17.6$ ), following the method presented in Sect. 4.1.

GRB 060607A occurred in a galaxy at  $z = 3.0749$ . It has the lowest neutral-hydrogen column density observed among the LGRB afterglow sample ( $\log(N_{\text{HI}}/\text{cm}^{-2}) = 16.95 \pm 0.03$ ; Tanvir et al. 2019) and, like GRB 191004B, the only metal absorption lines at the GRB redshift are those of C IV and Si IV. The VLT/UVES spectrum of its afterglow shows a very complex line of sight, rich in intervening absorbers (Fox et al. 2008; Prochaska et al. 2008; Fynbo et al. 2009). There is evidence of LyC emission over  $\sim 16 \text{ \AA}$  below the LL, likely absorbed bluewards from an intervening system. It is not firmly possible to exclude that the apparent LyC emission is due to an interloper. Nonetheless, deep *HST-F775W* imaging observations are available for the GRB 060607A field (see Blanchard et al. 2016) and no galaxy is detected at  $m_{AB} > 28.9$  mag over a radius of  $\sim 15$  kpc from the afterglow position. From the UVES spectrum, we measure the ionizing ( $f_{900}^{\text{obs}}$ ) flux density over the  $16 \text{ \AA}$  where LyC emission is detected and we find a  $(f_{1500}/f_{900})^{\text{obs}}$  ratio of  $5.0 \pm 0.6$ . The determination of dust extinction on this line of sight is not trivial as the afterglow experienced some rebrightening during the X-ray observations and UVES spectroscopy. Following the same prescriptions as in Sect. 3.3, we find a quite high extinction with  $A_V = 0.13 \pm 0.04$  mag for the Small Magellanic Cloud extinction curve (see also Kann et al. 2010). Following the same method as for the previous cases, we derive  $(f_{1500}/f_{900})^{\text{int}} = 0.42$ . The distribution of  $f_{\text{esc,rel}}(\text{LyC})$  is presented in Fig. 7 and shows that 95% of the distribution is at  $f_{\text{esc,abs}} \geq 0.05$  ( $f_{\text{esc,rel}} \geq 0.09$ ) and 96% with  $f_{\text{esc,rel}} \leq 1$ . The median escape fraction for the latter 96% is  $f_{\text{esc,abs}} = 0.08^{+0.03}_{-0.02}$  ( $f_{\text{esc,rel}} = 0.13^{+0.04}_{-0.03}$ ) with an average value of  $\langle f_{\text{esc,abs}} \rangle = 0.08$  ( $\langle f_{\text{esc,rel}} \rangle = 0.14$ ). The value derived from the HI column density, following the method presented in Sect. 4.1, is  $f_{\text{esc}} = 0.45 \pm 0.13$  and  $f_{\text{esc,abs}} = 0.20 \pm 0.05$ , with  $\log(N_{\text{HI}}/\text{cm}^{-2}) = 17.1 \pm 0.15$  (in agreement with the column density reported in the literature). The difference of the  $f_{\text{esc}}$  values determined by the two methods could be due to a higher  $T_{900}^{\text{IGM}}$  average from simulations, as this line of sight is particularly rich in absorbing systems. We stress that both  $f_{\text{esc,abs}}$  values are tentative due to the difficulties in determining the dust extinction.

The  $f_{\text{esc}}$  results for the three GRBs are summarized in Table 2.

### 5.1. Statistical estimation of the ionizing escape fraction in LGRB host galaxies

The HI column density of  $\log(N_{\text{HI}}/\text{cm}^{-2}) = 17.2 \pm 0.15$  derived for GRB 191004B corresponds to one of the lowest  $N_{\text{HI}}$  measured for an LGRB host galaxy. Adding this case to the sample studied in Tanvir et al. (2019) we re-evaluate the average escape fraction derived from the compilation of the 141 column densities available to-date. To do so, we use the formula described in Sect. 4.1 (Eq. (3)) averaged over the full sample and we find  $\langle f_{\text{esc}} \rangle = 0.007$  instead of 0.005 presented in Tanvir et al. (2019). Following again the method used in Chen et al. (2007) and Tanvir et al. (2019) we performed a bootstrap resampling by allowing replacement of the 141 sightlines in the above sample. From the  $10^6$   $N_{\text{HI}}$  distributions simulated with this method and taking into account scatter produced by the uncertainty on each data point, we find a 98% c.l upper limit of  $\langle f_{\text{esc}} \rangle < 0.020$ . This new estimation increases somewhat the result found in Tanvir et al.



**Fig. 8.** Best fit of the GRB 191004B Ly $\alpha$  line, observed in the X-shooter host-galaxy spectrum, with unconstrained shell models. In the top panel, the solid black line corresponds to the data with their error bars (grey), in dashed orange the best fit and in dotted blue the intrinsic Ly $\alpha$  emission predicted by the model. In the top left corner, we show the best-fit parameters. In the bottom panel, the dots correspond to the normalized residuals between the observation and the model. The orange ones are for the Ly $\alpha$  profile only with the corresponding p-value for the Shapiro-Wilk test.

(2019) but is still consistent with their  $\langle f_{\text{esc}} \rangle$  upper limit in the range 0.015 - 0.02.

## 6. Modelling of the Ly $\alpha$ emission line

The non detection of nebular emission lines in the spectra of GRB 191004B prevents firm determination of the exact redshift of the host galaxy. We can, however, use the empirical correlation found by Verhamme et al. (2018) between the FWHM and the velocity shift of the peak of the Ly $\alpha$  line. We derive a predicted redshift of  $z_{\text{predict}} = 3.5060 \pm 0.0008$ . This value is fully consistent with the HI component seen in absorption at  $z = 3.5055$ .

Due to its resonant nature, the Ly $\alpha$  radiation produced by active star-forming regions carries the signature of the physical properties of the neutral gas where the photons scatter before escaping. In that sense, the detection of the Ly $\alpha$  line gives us a unique chance to investigate if the properties probed by the GRB afterglow lines of sight are similar to the ones of the material probed by the Ly $\alpha$  emission. To do this comparison, we perform a *shell-model* fitting of the observed Ly $\alpha$  line using an improved model grid and process originally described in Gronke et al. (2015). The shell model is a commonly used model (first introduced by Ahn et al. 2003, also see Verhamme et al. 2006) that consists of a central source emitting a continuum and a Gaussian Ly $\alpha$  line defined by an intrinsic width and equivalent width ( $\sigma_i$ ,  $EW_i$ ). This source is surrounded by a shell of neutral hydrogen and described by four parameters: a radial expansion velocity  $v_{\text{exp}}$ ; a HI column density  $N_{\text{HI}}$ ; an effective temperature of the gas  $T$ ; an optical depth of dust close to Ly $\alpha$  wavelength  $\tau_d$ . We

fit the X-shooter host-galaxy data leaving these parameters free in the grid range values described in Gronke et al. (2015). As the final fitting result is sensitive to the redshift of the emitting source, we also leave the systemic redshift as a free parameter with a Gaussian prior of  $z_{\text{syst}} = 3.5060 \pm 0.01$ , based on the estimation derived above. Note, however, that the prior on  $z_{\text{syst}}$  is very wide, and thus, essentially a free parameter.

We compare the observed Ly $\alpha$  profile to the best-fit model in Fig. 8. We find that the best-fit profile describes the observation very well. Indeed, the normalized residuals between the observed and modeled Ly $\alpha$  line (brown dots in Fig. 8) provides a Shapiro-Wilk coefficient of 0.985 and a p-value = 0.316, therefore consistent with Gaussian noise. The best-fit results show a small blue component at  $-210 \text{ km s}^{-1}$ , consistent with the spectrum. This blue component is only detected at a  $2.5\sigma$  confidence level, with a flux of  $F(\text{Ly}\alpha, \text{blue}) = (0.5 \pm 0.2) \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ . If real, the separation between the blue and red peaks is  $323 \text{ km s}^{-1}$ .

The best-fit parameters are consistent with the observed values, (i): the systemic redshift  $z_{\text{syst}} = 3.5053^{+0.0005}_{-0.0003}$  is in good agreement with the value of the reddest HI absorption component seen at  $z = 3.5055$  and also with  $z_{\text{predict}} = 3.5060 \pm 0.0008$ ; (ii): the low expansion velocity of the gas  $v_{\text{exp}} = 40.0^{+12.0}_{-11.0} \text{ km s}^{-1}$  is consistent with the velocity of the reddest HI component which also is the richest in neutral hydrogen (cf. Fig. 2); (iii): the predicted HI column density,  $\log(N_{\text{HI}}/\text{cm}^{-2}) = 17.7^{+0.7}_{-0.7}$  and (iv): the dust optical depth  $\tau_d = 0.3^{+0.5}_{-0.2}$  corresponding to  $A_V = 0.04^{+0.07}_{-0.04} \text{ mag}$  are fully consistent with the observational values derived from the afterglow spectrum (see Sections 3.3 and 3.1); (v): the intrinsic properties of Ly $\alpha$  and Balmer lines are expected to be similar if we assume both transitions to be formed by recombination in the same regions of the galaxy. Here, we cannot constrain the intrinsic properties of the Ly $\alpha$  source ( $\sigma_i$ ,  $EW_i$ ) as we do not observe the Balmer lines. Nevertheless, by considering the recombination factor 8.7 (Sect. 3.4) and the H $\alpha$ /H $\beta$  flux ratio of 2.86 (Sect. 3.5), the intrinsic Ly $\alpha$  properties predicted by the best-fit model ( $\sigma_i = 164^{+29}_{-33} \text{ km s}^{-1}$ ,  $EW_i(\text{Ly}\alpha) = 24^{+13}_{-9} \text{ \AA}$ ) imply a rest-frame  $EW(\text{H}\beta) = 1.0^{+0.6}_{-0.4} \text{ \AA}$  which is fully consistent with its non detection in the X-shooter spectrum ( $EW(\text{H}\beta) < 3.0 \text{ \AA}$ , see Sect. 3.5).

We stress that, while it is well known that the *shell-model* can reproduce observed Ly $\alpha$  spectra extremely well (e.g., Gronke 2017), its clear oversimplification of realistic HI configurations in and around galaxies is a source of debate in the literature parameters (e.g., Gronke et al. 2017; Orlitová et al. 2018).

## 7. Discussion

### 7.1. Comparison with Ly $\alpha$ -emitter LGRBs (LAE-LGRBs)

The Ly $\alpha$  line profile of GRB 191004B is typical of what is observed for known LAE-LGRBs. It shows a main peak redshifted by  $113 \pm 30 \text{ km s}^{-1}$  which is just below the range of values ( $150\text{-}750 \text{ km s}^{-1}$ ) determined for the known LAE-LGRB host-galaxies in Milvang-Jensen et al. (2012). The low-velocity shift of the line associated with LyC leakage is consistent with the theoretical predictions of Verhamme et al. (2015). Indeed, they found that the classical asymmetric redshifted profile of the Ly $\alpha$  line should be shifted by less than  $150 \text{ km s}^{-1}$  in the case of LyC leakage. This was also observationally confirmed in Verhamme et al. (2017) for low-redshift LyC emitters. The host galaxy has a Ly $\alpha$  luminosity ( $L_{\text{Ly}\alpha} = (1.18 \pm 0.18) \times 10^{42}$

erg s<sup>-1</sup>) and a UV magnitude lower limit ( $M_{UV} > -19.2$  mag) found to be close to the median values of the LAE-LGRB host-galaxies sample, respectively,  $L_{Ly\alpha} = 1.7^{+3.5}_{-0.9} \times 10^{42}$  erg s<sup>-1</sup> (Fynbo et al. 2003; Jakobsson et al. 2005; Milvang-Jensen et al. 2012) and  $M_{UV} = -19.9^{+1.4}_{-1.0}$  mag (Tanvir et al. 2019). However, GRB 191004B is the only LAE-LGRB for which LyC leakage has been detected (over the eight LAE-LGRBs with a spectrum covering the spectral range beyond the Lyman limit) or inferred from the HI column density probed by the afterglow. It is in fact the only LAE-LGRB having  $\log(N_{HI}/\text{cm}^{-2}) < 18$ , over the twenty-two LAE-LGRBs with an afterglow spectrum covering the Ly $\alpha$  absorption spectral range. Indeed, LGRBs usually probe dense HI column densities. Among the 140 LGRB lines of sight of the sample of Tanvir et al. (2019) only two have  $\log(N_{HI}/\text{cm}^{-2}) < 18$ , namely GRB 050908 ( $\log(N_{HI}/\text{cm}^{-2})=17.6$ ) and GRB 060607A ( $\log(N_{HI}/\text{cm}^{-2})=16.95$ ). Consistently, together with GRB 191004B, they also are the only cases showing non-zero afterglow continuum emission below the Lyman limit. Intriguingly they do not show any trace of Ly $\alpha$  emission above a  $3\sigma$  flux upper limit of  $F(\text{Ly}\alpha) = 6.4 \times 10^{-18}$  erg s<sup>-1</sup> cm<sup>-2</sup> and  $F(\text{Ly}\alpha) = 7.3 \times 10^{-18}$  erg s<sup>-1</sup> cm<sup>-2</sup>, respectively (Milvang-Jensen et al. 2012). We should consider, however, that those are extremely faint galaxies. Even if no firm conclusion can be drawn from such a small sample, the fact that the highest escape fraction is obtained for GRB 191004B, which is also the only LAE, seems to support the idea that strong LyC emission is correlated with Ly $\alpha$  emission. Inversely, substantial Ly $\alpha$  emission is not necessarily associated with LyC leakage, as GRB 191004B is the only LyC leaker among eight LAE-LGRBs of comparison.

LyC and Ly $\alpha$  escape are expected to be somehow correlated (see, e.g., Verhamme et al. 2015, 2017; Dijkstra et al. 2016; Kakiichi & Gronke 2019) since the escape of both radiations is facilitated by low- $N_{HI}$  sightlines. Nevertheless, there exist physical phenomena depending mainly on the density of the HI gas and the presence of low density paths in the medium, that can boost the Ly $\alpha$  against LyC escape, or inversely (see, e.g., Ji et al. 2020, Sect. 3.2). Indeed, recent studies such as the one of Bian & Fan (2020) show that strong Ly $\alpha$  emission is not necessarily a good indicator of LyC leakage (see also e.g., Guaita et al. 2016; Grazian et al. 2017). Observationally, the Ly $\alpha$  EW does seem to correlate with LyC escape as shown by Steidel et al. (2018). There are no deep enough observations to put significant Ly $\alpha$  EW limits for GRB 050908 and GRB 060607A hosts, but in general, galaxies with comparably high LyC  $f_{esc}$  as GRB 191004B have higher Ly $\alpha$  EW<sub>0</sub> (e.g. Izotov et al. 2016, 2018a; Vanzella et al. 2016). A relation between the separation of the observed Ly $\alpha$  emission blue and red peaks and the LyC escape fraction has also been found by e.g. Verhamme et al. (2015, 2017) and Izotov et al. (2018b) for low-redshift LyC leaking LAEs. The separation of  $\sim 300$  km s<sup>-1</sup> found for GRB 191004B would correspond to  $f_{esc} \sim 10\%$ , thus, lower than the value we determined. In our case, we are probing the LyC leakage along the GRB line of sight. Therefore, it is possible that the average LyC leakage of the galaxy and the  $N_{HI}$  through which Ly $\alpha$  photons escape are different than those probed by the GRB afterglow. A higher average  $N_{HI}$  value, as the best value of the shell-model fitting ( $\log(N_{HI}/\text{cm}^{-2}) = 17.7$ ), would imply an  $f_{esc}(\text{LyC})$  of 6.9%, in better agreement with the Ly $\alpha$  EW<sub>0</sub> and peak separation relations. In this case, our results could support a medium with low density holes and the escape would not be homogeneous. This is a more realistic scenario, as the ISM in galaxies is not homogeneous. We would like to point out also that the re-

lations above have not been tested for faint, high-redshift galaxies as those studied in this paper. Steidel et al. (2018) explore brighter galaxies, whereas Fletcher et al. (2019) study systems with Ly $\alpha$  EW<sub>0</sub> > 15 Å.

## 7.2. Comparison with known LyC emitters

We collected absolute escape fraction measurements from the literature, especially from galaxies lying at  $z \gtrsim 3$ . The comparison between our LGRB measurements and those from the literature is presented in Fig. 9 as a function of galaxy UV luminosity.

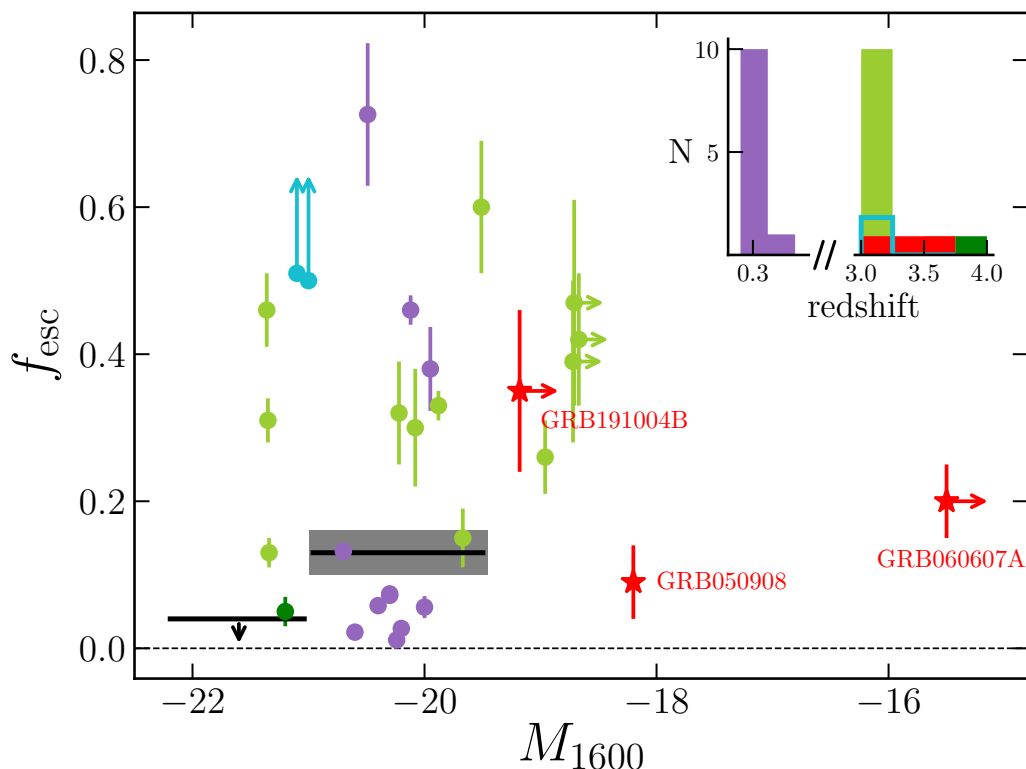
The data in the plot is not homogeneous, owing to the difficulty of finding LyC emitters and measuring their escape fractions. The IGM transmission  $T_{900}^{\text{IGM}}$  depends on the assumed model; Steidel et al. (2018) and Shapley et al. (2016) use a combined transmission of the IGM and circumgalactic medium (CGM), resulting in on average lower transmission than our model (and that of Vanzella et al. 2016). Fletcher et al. (2019) adopt an average transmission at  $z \sim 3$ . If they used a similar probabilistic approach as in this work, the measurements from their study would have lower values, and they would be only a lower limit. The choice of  $(f_{1500}/f_{900})^{\text{int}}$  also varies based on different galaxy population synthesis studies. Finally, we overplot the sample of  $z \sim 0.3$  green pea LyC emitters. Their low-redshift origin makes the measurements less affected by the uncertainties in  $T_{900}^{\text{IGM}}$  and  $(f_{1500}/f_{900})^{\text{int}}$ .

Fig. 9 shows that LGRBs can be powerful tools to probe LyC leakage from extremely faint high-redshift galaxies, making it possible to explore LyC leakage from galaxies beyond the current absolute magnitude limits of galaxy studies. The number of objects at  $M_{1600} > -19.5$  mag is similar to those of galaxy studies and the power to access even fainter galaxies is evident. LyC leakage investigations through LGRB afterglows do not suffer from galaxy apparent magnitude selection. Furthermore the results are obtained as a by-product of routinely performed observations for the GRB redshift determination, requiring a very small observing time ( $\sim 1$  - 3 hours).

We stress that there is a fundamental difference between our study, using LGRB afterglows, and the common ones using galaxy observations. In our case we look into the transparency of the galactic material to the ionizing radiation in one line of sight (toward the studied LGRB), while the other studies measure the averaged transparency of a whole galaxy facing us. Therefore through LGRB afterglow we can probe directly the lines of sight through which LyC photons escape.

The use of LGRBs to constrain the leakage of LyC photons can be a powerful way to understand what sources reionized the Universe. Indeed, LGRBs can probe  $f_{esc}$  not only at higher redshift – thus, requiring less extrapolation in terms of galactic properties – but also (UV) fainter galaxies. This is particularly noteworthy since those galaxies are thought to be the main driver of reionization because (i) models and analytical considerations show that their feedback is likely overruling gravity, thus, strong enough to clear channels through which LyC photons can escape (Paardekooper et al. 2015; Cen 2020), and (ii) the (increasing) steepness of the Ly $\alpha$  luminosity function hints toward a very steep LyC luminosity function (Gronke et al. 2015; Dressler et al. 2015; Dijkstra et al. 2016). A larger sample of LGRB afterglow spectra will allow us to test these theories.





**Fig. 9.** Collection of measurements of absolute escape fractions of known LyC emitters as a function of galaxy UV luminosity. The LGRB lines of sight studied in this paper are shown with red stars. Other data points are taken from Fletcher et al. (2019) ( $z \sim 3$ , light green), Izotov et al. (2016, 2018a,b) ( $z \sim 0.3$ , violet), Ji et al. (2020) ( $z \sim 3.8$ , dark green) and Vanzella et al. (2016) and Shapley et al. (2016) ( $z \sim 3.2$ , blue). Black lines indicate the limit and the average as provided by the  $z \sim 3$  sample study of Steidel et al. (2018). The inset shows the redshift distribution of the systems.

## 8. Conclusions

We investigated LyC leakage in LGRB afterglow spectra through the observations of GRB 191004B at  $z = 3.5055$  (presented here for the first time), GRB 050908 ( $z = 3.0749$ ), and GRB 060607A ( $z = 3.3467$ ). We found absolute escape fractions along their sightlines of  $0.35^{+0.10}_{-0.11}$ ,  $0.08^{+0.05}_{-0.04}$ ,  $0.20^{+0.05}_{-0.05}$ , respectively. By using simulations of the IGM opacity, we found similar results except for GRB 060607A likely because of its very rich line of sight.

Thanks to the fact that the intrinsic afterglow emission corresponds to a single power-law, the determination of the intrinsic flux density ratio between the fraction of escaping ionizing and non-ionizing photons does not suffer from the uncertainties due to the assumptions behind galaxy spectral-synthesis models. Furthermore afterglow-based studies do not suffer from galaxy magnitude selection effects as galaxy studies.

The results presented here are by-products of routinely performed observations to determine the redshift of GRBs obtained using only  $\sim 1$ -3 hours each, compared to the many hours needed for galaxy-based studies to pre-select and observe such faint galaxies. They show that LGRBs can be powerful tools to study LyC leakage from faint, star-forming galaxies at high redshift. Indeed, the host galaxies of the LGRBs presented here all have  $M_{1600} > -19.5$  mag, with the extreme case of the host of GRB 060607A at  $M_{1600} > -16$  mag. Such faint galaxies are very common at very high redshift. Their role in reionization is still debated, but their global ionizing photon budget may contribute significantly to the reionization process.

LGRBs explode in young star-forming regions, major contributors of the LyC photons inside galaxies. LGRB afterglows

probe lines of sight originating from these regions, therefore they shed light on the paths through which LyC photons escape. Uniquely LGRBs also have the potential to combine the study of these lines of sight (through afterglow spectroscopy) with those of the global properties of their hosting galaxies, through photometric and spectroscopic observations once the afterglow has disappeared. The galaxies presented in this paper are faint and at the limit of current instrumentation, but they will be perfect targets for E-ELT and JWST. Future observations will allow us to determine their characteristics and to define indirect indicators to be used to find similar galaxies at the reionization epoch.

The current limitation of the LGRB contribution to reionization studies is the limited number of LGRBs with afterglow detections, especially at high redshift. Future satellites such as *Gamow Explorer* (White 2020) and *THESEUS*<sup>2</sup> (pre-selected as M5 European Space Agency mission), will largely improve this situation. Indeed the latter will allow the detection and redshift estimate of about 150 GRBs at  $3 < z < 4$ , during its 3-year nominal duration (Amati et al. 2018; Ghirlanda et al. 2015). Furthermore, its detection of  $\sim 100$  LGRBs at  $z > 6$ , combined with space and ground based observations will allow us to directly study the galaxies contributing to reionization.

**Acknowledgements.** This work is part of the *BEaPro* project funded by the French *Agence Nationale de la Recherche* (ANR-16-CE31-0003). We thank Giancarlo Ghirlanda for providing useful information. MG was supported by NASA through the NASA Hubble Fellowship grant HST-HF2-51409 and acknowledges support from HST grants HST-GO-15643.017-A, HST-AR-15039.003-A, and XSEDE grant TG-AST180036. DAK acknowledges support from Spanish research project RTI2018-098104-J-I00 (GRBPhot). The Pan-

<sup>2</sup> <https://www.isdc.unige.ch/theseus/>

STARRS1 Surveys (PS1) and the PS1 public science archive have been made possible through contributions by the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, the Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under Grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation Grant No. AST-1238877, the University of Maryland, Eotvos Lorand University (ELTE), the Los Alamos National Laboratory, and the Gordon and Betty Moore Foundation. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. The NumPy (van der Walt et al. 2011), SciPy (Virtanen et al. 2020) and matplotlib (Hunter 2007) packages have been extensively used for the preparation and presentation of this work. The Cosmic DAWN center is funded by the DNRF. JPUF thanks the Carlsberg foundation for support.

## References

- Ahn, S.-H., Lee, H.-W., & Lee, H. M. 2003, *MNRAS*, 340, 863
- Amati, L., O'Brien, P., Götz, D., et al. 2018, *Advances in Space Research*, 62, 191
- Amorín, R., Grazian, A., Castellano, M., et al. 2014, *ApJ*, 788, L4
- Appenzeller, I., Fricke, K., Fürtig, W., et al. 1998, *The Messenger*, 94, 1
- Bergvall, N., Zackrisson, E., Andersson, B. G., et al. 2006, *A&A*, 448, 513
- Bian, F. & Fan, X. 2020, *MNRAS*, 493, L65
- Bian, F., Fan, X., McGreer, I., Cai, Z., & Jiang, L. 2017, *ApJ*, 837, L12
- Blanchard, P. K., Berger, E., & fai Fong, W. 2016, *ApJ*, 817, 144
- Bolmer, J., Ledoux, C., Wiseman, P., et al. 2019, *A&A*, 623, A43
- Borthakur, S., Heckman, T. M., Leitherer, C., & Overzier, R. A. 2014, *Science*, 346, 216
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, *ApJ*, 803, 34
- Brocklehurst, M. 1971, *MNRAS*, 153, 471
- Bruzual, G. & Charlot, S. 2003, *MNRAS*, 344, 1000
- Cano, Z., Izzo, L., de Ugarte Postigo, A., et al. 2017, *A&A*, 605, A107
- Cen, R. 2020, *ApJ*, 889, L22
- Cenko, S. B., Barthelmy, S. D., Beardmore, A. P., et al. 2019, *GCN*, 25947, 1
- Cenko, S. B., Fox, D. B., Penprase, B. E., et al. 2008, *ApJ*, 677, 441
- Chabrier, G. 2003, *PASP*, 115, 763
- Chen, H., Prochaska, J. X., Ramirez-Ruiz, E., et al. 2007, *ApJ*, 663, 420
- Chen, H.-W., Prochaska, J. X., & Gnedin, N. Y. 2007, *ApJ*, 667, L125
- Christensen, L., Fynbo, J. P. U., & Prochaska, J. X. 2011, *ApJ*, 19
- Claeysens, A., Richard, J., Blaizot, J., et al. 2019, *MNRAS*, 489, 5022
- Cooper, T. J., Simcoe, R. A., Cooksey, K. L., O'Meara, J. M., & Torrey, P. 2015, *ApJ*, 812, 58
- Cristiani, S., Serrano, L. M., Fontanot, F., Vanzella, E., & Monaco, P. 2016, *MNRAS*, 462, 2478
- De Cia, A., Starling, R. L. C., Wiersema, K., et al. 2011, *MNRAS*, 418, 129
- De Pasquale, M., Piro, L., Perna, R., et al. 2003, *ApJ*, 592, 1018
- de Ugarte Postigo, A., Fynbo, J. P. U., Thöne, C. C., et al. 2012, *A&A*, 548, A11
- D'Elia, V., Fynbo, J. P. U., Izzo, L., et al. 2019, *GCN*, 25978, 1
- Dijkstra, M., Gronke, M., & Venkatesan, A. 2016, *ApJ*, 828, 71
- Dressler, A., Henry, A., Martin, C. L., et al. 2015, *ApJ*, 806, 19
- Fletcher, T. J., Tang, M., Robertson, B. E., et al. 2019, *ApJ*, 878, 87
- Fontanot, F., Cristiani, S., Pfrommer, C., Cupani, G., & Vanzella, E. 2014, *MNRAS*, 438, 2097
- Fontanot, F., Cristiani, S., & Vanzella, E. 2012, *MNRAS*, 425, 1413
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306–312
- Fox, A. J., Ledoux, C., Vreeswijk, P. M., Smette, A., & Jaunsen, A. O. 2008, *A&A*, 491, 189
- Fox, A. J., Lehner, N., Tumlinson, J., et al. 2013, *ApJ*, 778, 187
- Fynbo, J. P. U., Jakobsson, P., Møller, P., et al. 2003, *A&A*, 406, L63
- Fynbo, J. P. U., Jakobsson, P., Prochaska, J. X., et al. 2009, *ApJ*, 185, 526
- Ghirlanda, G., Salvaterra, R., Ghisellini, G., et al. 2015, *MNRAS*, 448, 2514
- Grazian, A., Giallongo, E., Paris, D., et al. 2017, *A&A*, 602, A18
- Gronke, M. 2017, *A&A*, 608, A139
- Gronke, M., Bull, P., & Dijkstra, M. 2015, *ApJ*, 812, 123
- Gronke, M., Dijkstra, M., McCourt, M., & Peng Oh, S. 2017, *A&A*, 607, A71
- Gronke, M., Dijkstra, M., Trenti, M., & Wyithe, S. 2015, *MNRAS*, 449, 1284
- Guaita, L., Pentericci, L., Grazian, A., et al. 2016, *A&A*, 587, A133
- Heintz, K. E., Watson, D., Jakobsson, P., et al. 2018, *MNRAS*, 479, 3456
- Hjorth, J., Malesani, D., Jakobsson, P., et al. 2012, *ApJ*, 756, 187
- Hjorth, J., Sollerman, J., Møller, P., et al. 2003, *Nature*, 423, 847
- Hunter, J. D. 2007, *Computing in Science & Engineering*, 9, 90
- Inoue, A. K., Iwata, I., Deharveng, J.-M., Buat, V., & Burgarella, D. 2005, *A&A*, 435, 471
- Izotov, Y. I., Schaerer, D., Thuan, T. X., et al. 2016, *MNRAS*, 461, 3683
- Izotov, Y. I., Schaerer, D., Worseck, G., et al. 2018a, *MNRAS*, 474, 4514
- Izotov, Y. I., Worseck, G., Schaerer, D., et al. 2018b, *MNRAS*, 478, 4851
- Jakobsson, P., Björnsson, G., Fynbo, J. P. U., et al. 2005, *MNRAS*, 362, 245
- Japelj, J., Covino, S., Gomboc, A., et al. 2015, *A&A*, 579, A74
- Japelj, J., Vanzella, E., Fontanot, F., et al. 2017, *MNRAS*, 468, 389
- Ji, Z., Giallisco, M., Vanzella, E., et al. 2020, *ApJ*, 888, 109
- Kakiichi, K. & Gronke, M. 2019, arXiv e-prints, arXiv:1905.02480
- Kann, D. A., Klose, S., Zhang, B., et al. 2010, *ApJ*, 720, 1513
- Kennicutt, Robert C., J. 1998, *ARA&A*, 36, 189
- Krogager, J.-K. 2018, arXiv e-prints, arXiv:1803.01187
- Kumar, P. & Zhang, B. 2015, *Phys. Rep.*, 561, 1
- LaPorte, S. J. & Cenke, S. B. 2019, *GCN*, 25978, 1
- Leitet, E., Bergvall, N., Hayes, M., Linné, S., & Zackrisson, E. 2013, *A&A*, 553, A106
- Madau, P. 1995, *ApJ*, 441, 18
- Matthee, J., Sobral, D., Gronke, M., et al. 2019, *MNRAS*, 492, 1778–1790
- Milvang-Jensen, B., Fynbo, J. P. U., Malesani, D., et al. 2012, *ApJ*, 756, 25
- Modigliani, A., Goldoni, P., Royer, F., et al. 2010, *SPIE Conference Series*, Vol. 7737, The X-shooter pipeline, 773728
- Nakajima, K., Ellis, R. S., Iwata, I., et al. 2016, *ApJ*, 831, L9
- Nakajima, K. & Ouchi, M. 2014, *MNRAS*, 442, 900
- Orlitová, I., Verhamme, A., Henry, A., et al. 2018, *A&A*, 616, A60
- Osterbrock, D. E. 1989, *Astrophysics of gaseous nebulae and active galactic nuclei* (ApJ)
- Paardekooper, J.-P., Khochfar, S., & Dalla Vecchia, C. 2015, *MNRAS*, 451, 2544
- Palmerio, J. T., Vergani, S. D., Salvaterra, R., et al. 2019, *A&A*, 623, A26
- Pei, Y. C. 1992, *ApJ*, 395, 130
- Perley, D. A., Tanvir, N. R., Hjorth, J., et al. 2016, *ApJ*, 817, 8
- Péroux, C., McMahon, R. G., Storrie-Lombardi, L. J., & Irwin, M. J. 2003, *MNRAS*, 346, 1103
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, *A&A*, 594, A13
- Prochaska, J. X., Dessauges-Zavadsky, M., Ramirez-Ruiz, E., & Chen, H.-W. 2008, *ApJ*, 685, 344
- Prochaska, J. X., O'Meara, J. M., & Worseck, G. 2010, *ApJ*, 718, 392
- Ramachandran, V., Hamann, W. R., Hainich, R., et al. 2018, *A&A*, 615, A40
- Rauch, M. 1998, *ARA&A*, 36, 267
- Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, *ApJ*, 802, L19
- Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, *ApJ*, 768, 71
- Salvaterra, R., Valle, M. D., Campana, S., et al. 2009, *Nature*, 461, 1258
- Schlafly, E. F. & Finkbeiner, D. P. 2011, *ApJ*, 737, 103
- Schulze, S., Fynbo, J. P. U., Milvang-Jensen, B., et al. 2012, *A&A*, 546, A20
- Selsing, J., Malesani, D., Goldoni, P., et al. 2019, *A&A*, 623, A92
- Shapley, A. E., Steidel, C. C., Strom, A. L., et al. 2016, *ApJ*, 826, L24
- Siana, B., Teplitz, H. I., Colbert, J., et al. 2007, *ApJ*, 668, 62
- Stamatikos, M., Barthelmy, S. D., Cenko, S. D., et al. 2019, *GCN*, 25955, 1
- Steidel, C. C., Bogosavljević, M., Shapley, A. E., et al. 2018, *ApJ*, 869, 123
- Steidel, C. C., Pettini, M., & Adelberger, K. L. 2001, *ApJ*, 546, 665
- Tanvir, N. R., Fox, D. B., Levan, A. J., et al. 2009, *Nature*, 461, 1254
- Tanvir, N. R., Fynbo, J. P. U., de Ugarte Postigo, A., et al. 2019, *MNRAS*, 483, 5380
- Tanvir, N. R., Levan, A. J., Fruchter, A. S., et al. 2012, *ApJ*, 754, 46
- Thöne, C. C., Christensen, L., Prochaska, J. X., et al. 2014, *MNRAS*, 441, 2034
- Thöne, C. C., Campana, S., Lazzati, D., et al. 2011, *MNRAS*, 414, 479
- Updike, A. C., Haislip, J. B., Nysewander, M. C., et al. 2008, *ApJ*, 685, 361
- van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *Computing in Science Engineering*, 13, 22
- van Dokkum, P. G. 2001, *PASP*, 113, 1420
- Vanzella, E., de Barros, S., Castellano, M., et al. 2015, *A&A*, 576, A116
- Vanzella, E., de Barros, S., Vasei, K., et al. 2016, *ApJ*, 825, 41
- Vanzella, E., Nonino, M., Cupani, G., et al. 2018, *MNRAS*, 476, L15
- Vergani, S. D. 2008, PhD thesis, School of Physical Sciences, Dublin City University
- Vergani, S. D., Salvaterra, R., Japelj, J., et al. 2015, *A&A*, 581, A102
- Verhamme, A., Garel, T., Ventou, E., et al. 2018, *MNRAS*, 478, L60
- Verhamme, A., Orlitová, I., Schaerer, D., & Hayes, M. 2015, *A&A*, 578, A7
- Verhamme, A., Orlitová, I., Schaerer, D., et al. 2017, *A&A*, 597, A13
- Verhamme, A., Schaerer, D., & Maselli, A. 2006, *A&A*, 460, 397
- Vernet, J., Dekker, H., D'Odorico, S., et al. 2011, *A&A*, 536, A105
- Vernet, J., Kerber, F., Mainieri, V., et al. 2009, *Proc. IAU*, 5, 535
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature Methods*, 17, 261
- White, N. E. 2020, arXiv e-prints, arXiv:2003.01592
- Woosley, S. E. & Heger, A. 2006, *ApJ*, 637, 914

Zafar, T., Watson, D., Møller, P., et al. 2018, MNRAS, 479, 1542

- 
- <sup>1</sup> GEPI, Observatoire de Paris, PSL Research University, CNRS, Place Jules Janssen, 92190 Meudon, France
- <sup>2</sup> Institut d'Astrophysique de Paris, UMR 7095, CNRS-SU, 98 bis boulevard Arago, 75014, Paris, France
- <sup>3</sup> Anton Pannekoek Institute for Astronomy, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
- <sup>4</sup> Cosmic Dawn Center (DAWN)
- <sup>5</sup> Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen Ø, Denmark
- <sup>6</sup> Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA
- <sup>7</sup> Centre for Astrophysics and Cosmology, Science Institute, University of Iceland, Dunhagi 5, 107 Reykjavík, Iceland
- <sup>8</sup> Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen Ø, Denmark
- <sup>9</sup> Department of Physics Astronomy and Leicester Institute of Space Earth Observation, University of Leicester, University Road, Leicester LE1 7RH, UK
- <sup>10</sup> ASI-Space Science Data Center, via del Politecnico snc, 00133 Rome, Italy
- <sup>11</sup> INAF – Osservatorio Astronomico di Roma, via Frascati 33, I-00040 Monteporzio Catone, Italy
- <sup>12</sup> Instituto de Astrofísica de Andalucía (IAA-CSIC), Glorieta de la Astronomía s/n, 18008 Granada, Spain
- <sup>13</sup> INAF – IASF/Milano, via Corti 12, I-20133 Milano, Italy
- <sup>14</sup> Department of Physics and Astronomy, University of Leicester, University Road, Leicester LE1 7RH, UK
- <sup>15</sup> Department of Physics, University of Warwick, Coventry CV4 7AL, UK
- <sup>16</sup> IRFU, CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France
- <sup>17</sup> Université Paris Diderot, AIM, Sorbonne Paris Cité, CEA, CNRS, 91191 Gif-sur-Yvette, France
- <sup>18</sup> INAF - Osservatorio Astronomico di Brera, via Bianchi 46, 23807, Merate (LC), Italy
- <sup>19</sup> Department of Astronomy and Space Sciences, Istanbul University, Beyazıt, 34119, Istanbul, Turkey
- <sup>20</sup> Department of Physics and Astronomy, Clemson University, Clemson, SC29634-0978, USA
- <sup>21</sup> Department of Physics, the George Washington University, 725 21st Street NW, Washington, DC 20052, USA
- <sup>22</sup> Astronomy, Physics and Statistics Institute of Sciences (APSIS), 725 21st Street NW, Washington, DC 20052, USA
- <sup>23</sup> Department of Physics, University of Bath, Claverton Down, Bath, BA2 7AY, UK
- <sup>24</sup> Department of Astrophysics/IMAPP, Radboud University, Nijmegen, The Netherlands
- <sup>25</sup> INAF – Osservatorio di Astrofisica e Scienza dello Spazio, via Piero Gobetti 93/3, 40129 Bologna, Italy